

# Report Assignment – 1

MT18086 (Vikash Kumar Pandey)

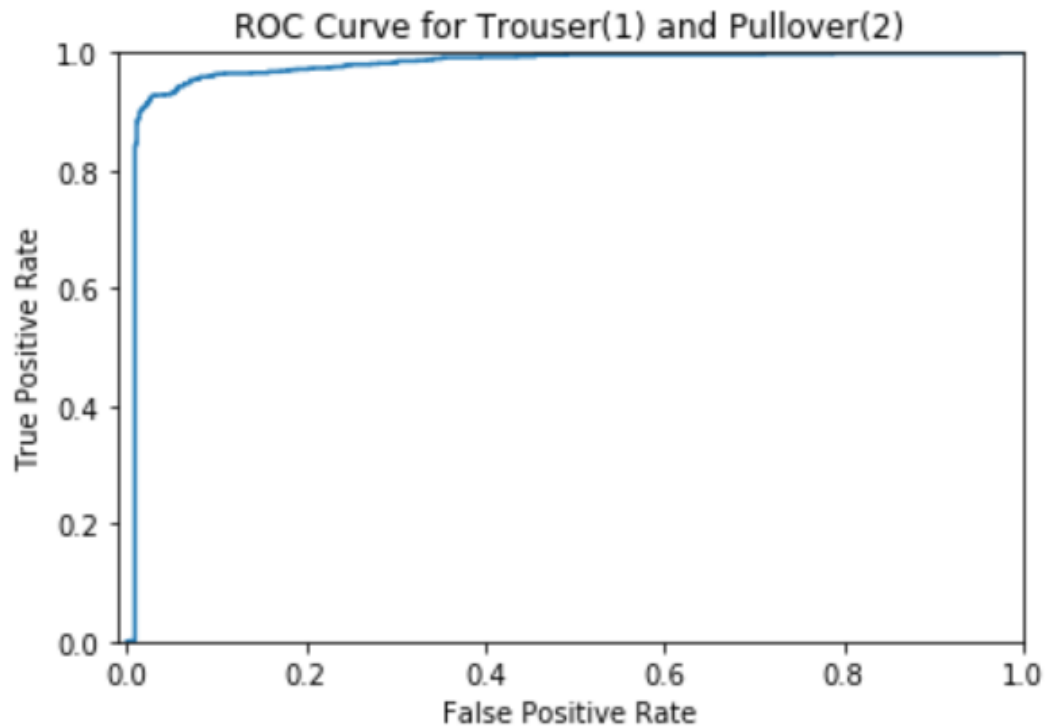
## Question 1-

### Part 1-

I - Confusion Matrix for two classes: Trouser and Pullover.

Actual Label	Predicted Label	
	T(Trouser)	F(Pullover)
	T(Trouser)	964
F(Pullover)	95	905

II -ROC Curve for two classes: Trouser and Pullover.



III -Precision, Recall for two classes: Trouser and Pullover.

Precision = 0.91029272899

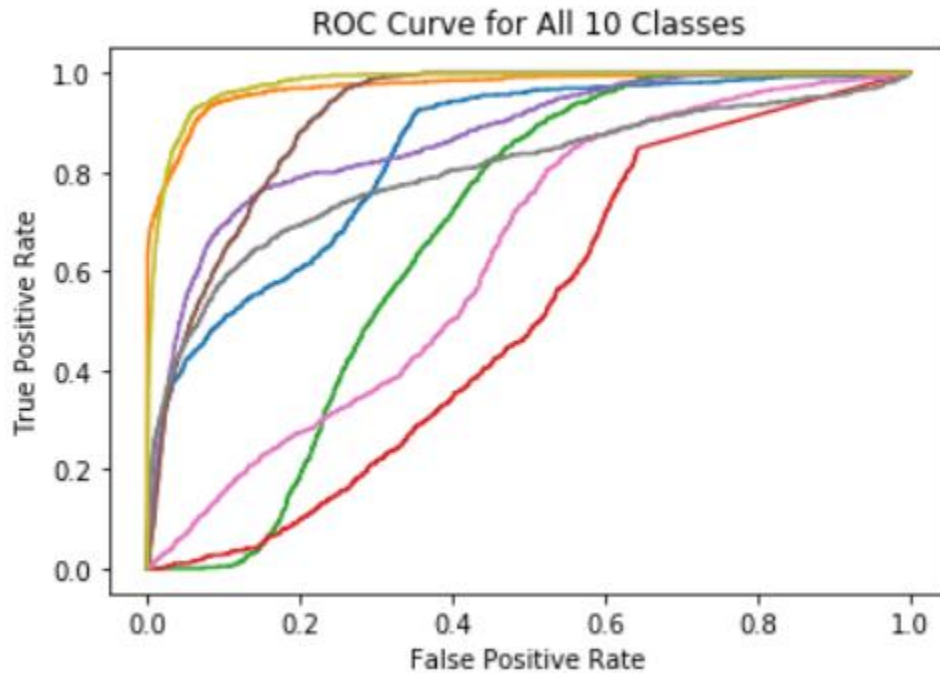
Recall = 0.964

**Part2-**

**I - Confusion Matrix:**

	Predicted Label										
TRUE LABEL		0_PR	1_PR	2_PR	3_PR	4_PR	5_PR	6_PR	7_PR	8_PR	9_Pr
	0_Act	603	31	25	92	29	105	103	0	12	0
	1_Act	27	878	3	54	14	12	10	0	2	0
	2_Act	4	6	281	11	352	106	218	0	22	0
	3_Act	34	17	1	732	67	95	50	0	4	0
	4_Act	2	3	62	64	708	63	87	0	11	0
	5_Act	0	0	0	1	0	733	7	187	5	67
	6_Act	169	4	76	55	271	176	212	0	37	0
	7_Act	0	0	0	0	0	130	0	805	0	65
	8_Act	3	1	9	45	11	72	61	9	788	1
	9_Act	0	0	0	1	0	67	12	59	3	858

## II - ROC Curve



## III - Precision and Recall for classes 1-10

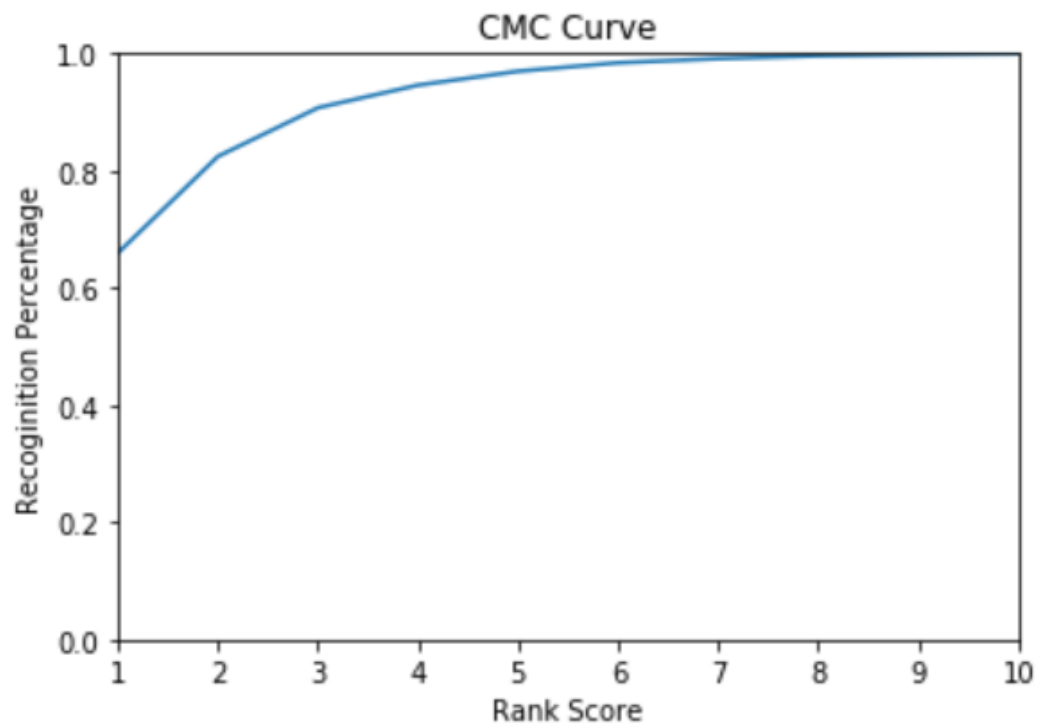
**Precision** = [0.7161520190023754, 1.6501945721938647, 2.2650742220844555, 2.958910846436973, 3.446516390428821, 3.916689578369809, 4.195636946790861, 4.955070909055012, 5.846473623987139, 6.712265753149601]

**Avg. Precision** = 67.12265753149602 %

**Recall** = [0.603, 1.4809999999999999, 1.762, 2.4939999999999998, 3.202, 3.935, 4.147, 4.952, 5.74, 6.598]

**Avg. Recall** = 65.97999999999999 %

**IV – CMC curve for class (1 to 10)**



## Question 2-

1 – A) In this part I apply Gaussian distribution without Binarization and get

Accuracy = 94.1204362257

Precision = 0.935400516796

Recall = 0.956828193833

Actual Label	Predicted Label		
		T(class=1)	F(class=8)
	T(class=1)	1086	49
	F(class=8)	75	899

B) Random 5-fold with cross validation for class 1 and 8

Folds	Accuracy
1	90.6312028583%
2	90.8693926161%
3	90.4724096864%
4	91.2266772529%
5	91.4580850219%

Mean Accuracy = 90.93155348710152%

Standard Deviation = 0.40903

Stratified 5-fold with cross validation for class 1 and 8

Folds	Accuracy
1	90.7899960302%
2	90.9090909091%
3	90.2342199285%
4	91.2663755459%
5	91.5772745332%

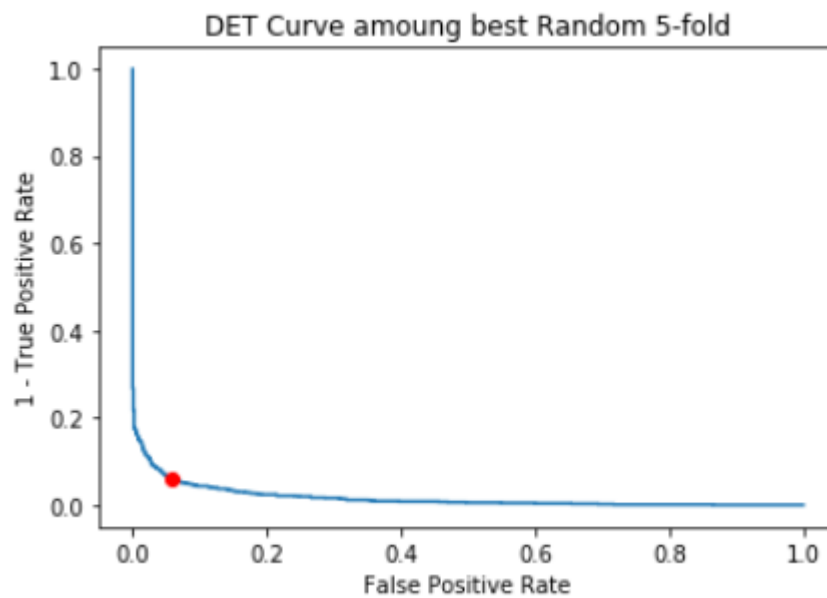
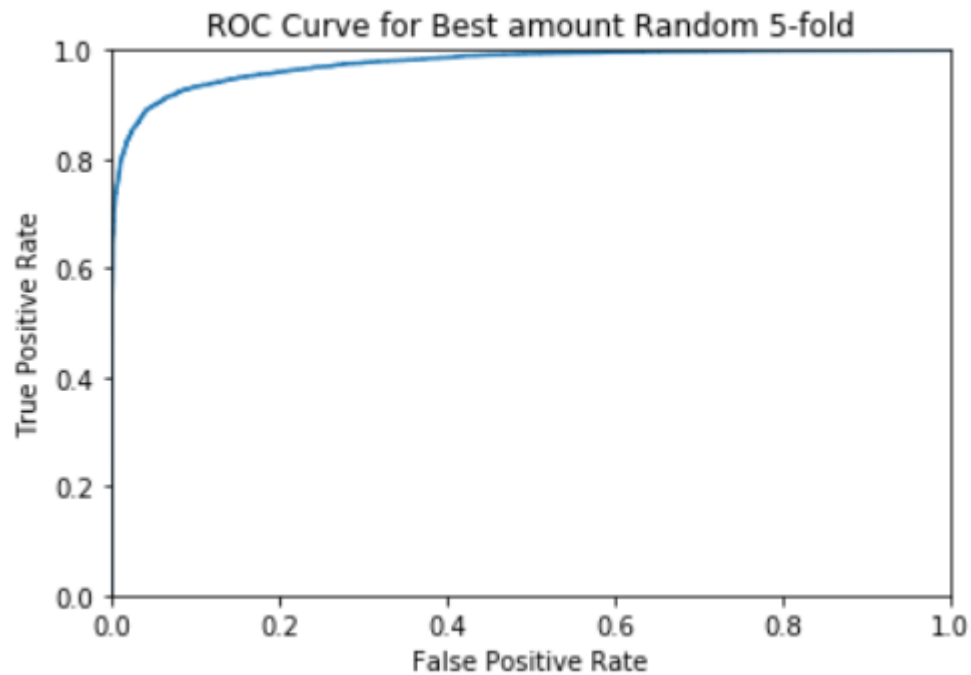
Average Accuracy = 90.95539138936613%

Standard Deviation = 0.50819

Because the standard deviation of Random-Cross fold is less than the stratified 5-fold so the best model chosen as Random 5-foldcross validation.

**C) I) DET and ROC Curves**

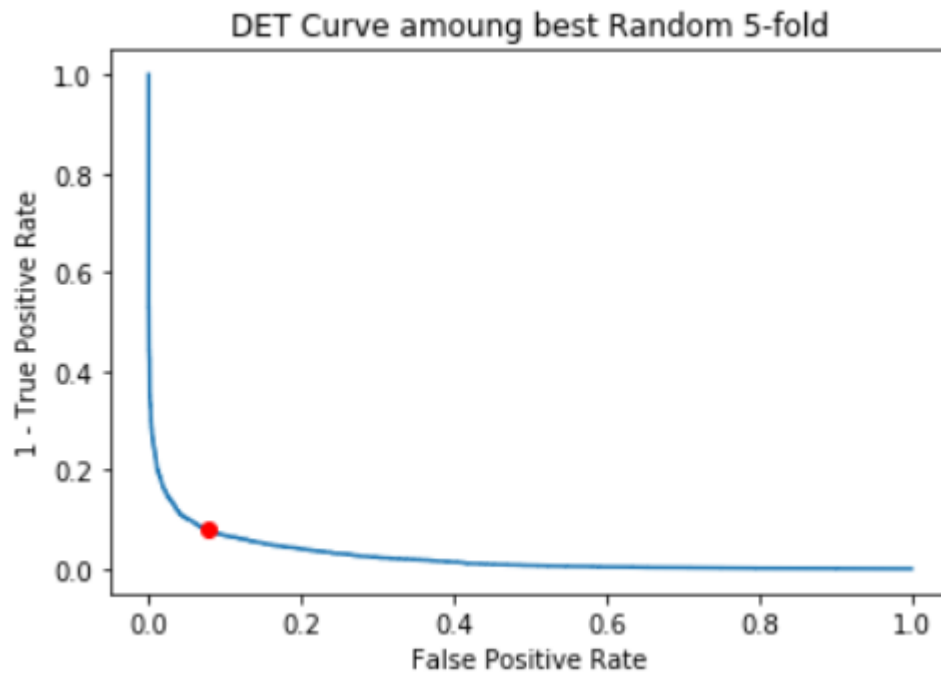
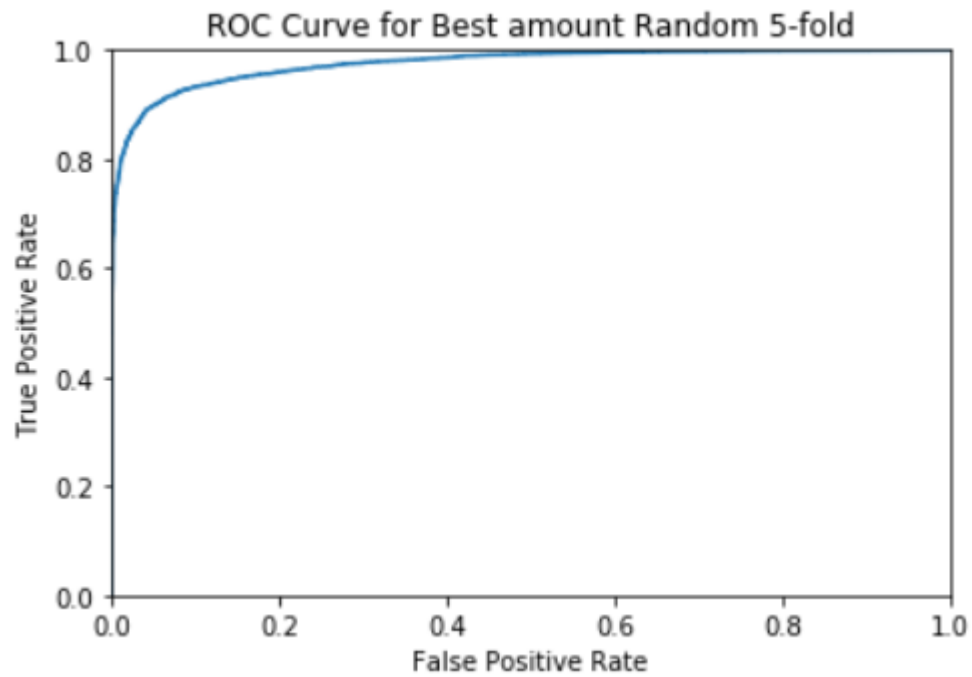
ROC and DET from best train model among 5-Fold for **Test data**.



Equal Error Rate = (0.058521560574948665, 0.058149779735682805)

For **False negative rate** = 0.058521560574948665, and **FPR** = 0.058149779735682805 value is Minimum, so the point define the EER.

ROC and DET from best train model among 5-Fold for **Training data**.



Equal Error Rate = (0.07810630661425397, 0.07816671610797987)

For **False negative rate** = 0.07810630661425397, and **FPR** = 0.07816671610797987 value is Minimum, so the point define the EER.

## II) Confusion matrix

For training data -

Actual Label	Predicted Label		
		T(class=1)	F(class=8)
	T(class=1)	6453	289
	F(class=8)	828	5023

For testing data –

Actual Label	Predicted Label		
		T(class=1)	F(class=8)
	T(class=1)	1080	55
	F(class=8)	84	890

2) A)

Accuracy = 69.8588709677%

Precision = 0.949781659389%

Recall = 0.430693069307%

Actual Label	Predicted Label		
		T(class=3)	F(class=8)
	T(class=3)	435	575
	F(class=8)	23	951

B) Random 5-fold with cross validation for class 3 and 8

Folds	Accuracy
1	71.5060492282%
2	73.9674593242%
3	69.5035460993%
4	68.6274509804%
5	69.6324143693%



Average Accuracy = 70.64738400025931%  
Standard Deviation = 2.13183

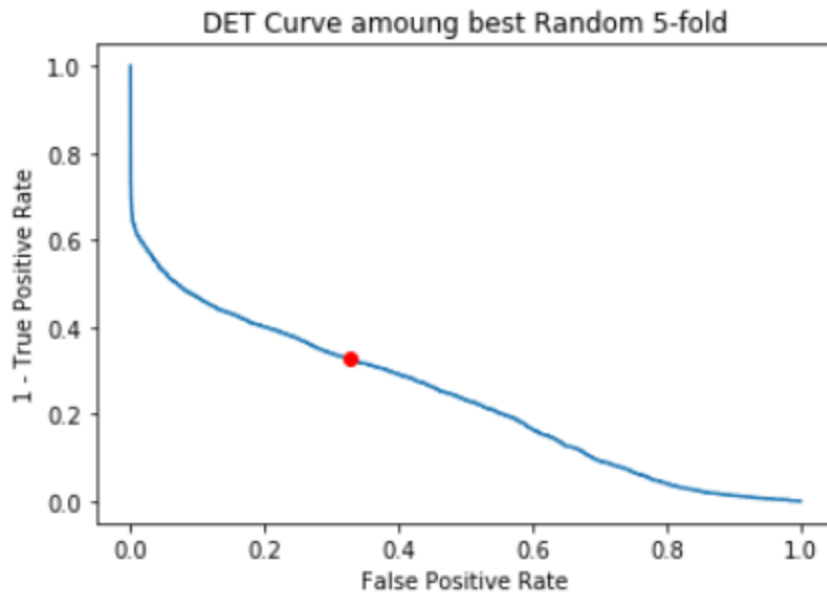
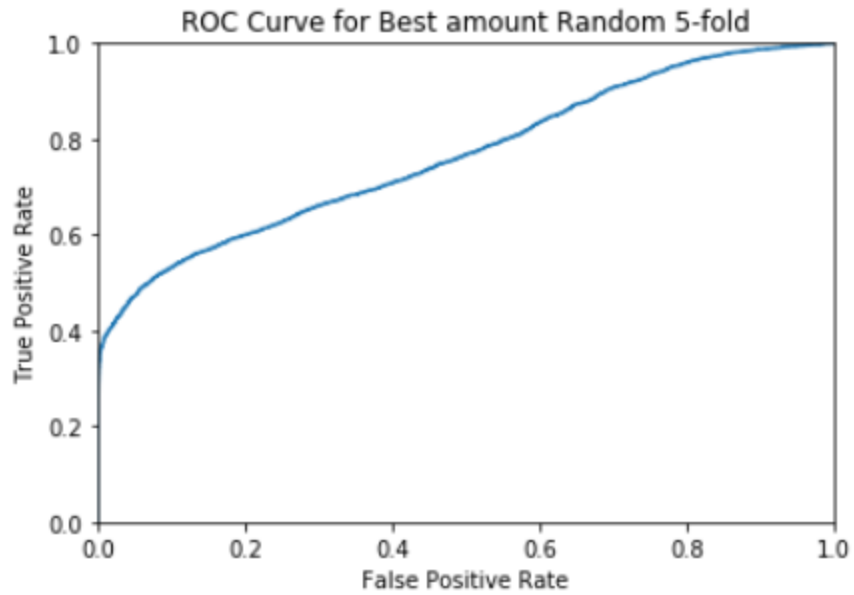
**Stratified 5-fold with cross validation for class 3 and 8**

Folds	Accuracy
1	70.5465164789%
2	71.0471422612%
3	70.1293283271%
4	70.4213600334%
5	74.1436925647%

Average Accuracy = 71.257608%  
Standard Deviation = 1.6471

**C) DET and ROC Curves**

Accuracy is approximate same but the standard deviation of stratified 5-fold is less than the Random 5-fold so the best model chosen as Stratified 5-fold cross validation. For training – Because the shape of the both 3 and 8 is quite similar so its difficult for classifier to detect 3, 8 that's why the accuracy is not good.



Equal Error Rate = (0.3271235686207486, 0.3271896917305497)

For **False negative rate** = 0.3271235686207486, and **FPR** = 0.3271896917305497 value is Minimum, so the point define the EER.

For testing

Actual Label	Predicted Label		
		T(class=1)	F(class=8)
	T(class=1)	6453	289
	F(class=8)	820	5031

